V. Vedral
Optics Section, Blackett Laboratory, Imperial College, Prince Consort Road, London SW7 2BZ
(July 22, 2002)

We analyze a quantum measurement where the apparatus is initially in a mixed state. We show that the amount of information gained in a measurement is not equal to the amount of entanglement between the system and the apparatus, but is instead equal to the degree of classical correlations between the two. As a consequence, we derive an uncertainty-like expression relating the information gain in the measurement and the initial mixedness of the apparatus. Final entanglement between the environment and the apparatus is also shown to be relevant for the efficiency of the measurement.

Any measurement can be modeled as an establishment of correlations between two random variables: one random variable represents the values of the quantity pertaining to the system to be measured, while the other random variable represents the states of the apparatus used to measure the systems [1]. It is by looking at the states of the apparatus, and discriminating them, that we infer the states of the system. Looking at the apparatus, of course, is another measurement process itself, which correlates our mental states (presumably another random variable) with those of the apparatus, so that indirectly we become correlated with the system as well. It is at this point that we can say that we have gained a certain amount of information about the system. This description of the measurement process is true both in classical and quantum physics. (Note that this way there is no more mystery in the "quantum state collapse" than there is in the corresponding classical measurement). The difference between the two lies in the way we represent states of systems and the way we represent their mutual interaction and evolution. Classically, physical states of an n-dimensional system are vectors in a real n dimensional vector space whose elements are various occupational probabilities for the states. The evolution of a classical system is in general some stochastic map acting on this vector space. Quantum mechanically, on the other hand, states are in general represented using density matrices, while the evolution is a completely positive, trace preserving transformation acting on these matrices. Using this representation, classical physics becomes a limiting case of quantum mechanics when the density matrices are strictly diagonal in one and the same fixed basis and the completely positive map then becomes the stochastic map. Because of this fact, it is enough to analyze properties of quantum systems and quantum evolutions and all the results are automatically applicable to classical physics when we restrict ourselves to the diagonal density operators only. A comprehensive survey of major papers on quantum measurement can be found in [2] and the first fully quantum analysis was due to von Neumann [3].

In this letter we analyze a quantum measurement when the apparatus is "fuzzy", i.e. it is initially in a mixed state. Our approach is entropic in character and is therefore closest in spirit to that of Lindblad [4]. We show that the amount of information gained via the apparatus is proportional to the *classical* correlations between the systems and the apparatus, rather than the amount of entanglement between them. We then derive an uncertainty-like expression which says that the sum of the information gained in the measurement and the mixedness of the apparatus (quantified by the von Neumann entropy [3]) is bounded from the above by $\log N$, where N is the dimension of the apparatus. Our analysis builds on recent results in quantum information theory concerning quantification of entanglement in bi- [5,6] and tripartite systems [7] and separating classical and quantum correlations [8]. Quantum information theory has mainly been developed to understand computation and communication supported by quantum systems, but this knowledge can now be applied back to quantum mechanics to study its foundations from a new perspective.

In quantum information theory it is common to distinguish between purely classical information, measured in bits, and quantum information, which is measured in qubits [9]. These differ in the channel resources required to communicate them. Qubits cannot be sent by a classical channel alone, but must be sent either via a quantum channel which preserves coherence or by teleportation through an entangled channel with two classical bits of communication. In this context, one qubit is equivalent to one unit of shared entanglement, or 'e-bit', together with two classical bits. Any bipartite quantum state may be used as a communication channel, and so it is of interest to determine how to separate the correlations it contains into a classical and an entangled part. A number of measures of entanglement and of total correlations have been proposed in recent years [9]. Recently we have also suggested how to quantify the purely classical part of the total bipartite correlations [8]. It is this quantity that will represent the information gain in a quantum measurement.

---

*This work is dedicated to Michael Vincent Vedral.

We first review the existing measures of entangled and total correlations. In classical information theory, the Shannon entropy, $H(X) \equiv H(p) = -\sum_i p_i \log p_i$, is used to quantify the information in a random variable, X, that contains states x_i with probabilities p_i [10]. The relative entropy is a useful measure of the closeness of two probability distributions $\{p_i\}$ and $\{q_i\}$ from the same random process X. The relative entropy of $\{p_i\}$ to $\{q_i\}$ is defined as $H(p||q) = \sum_i p_i \log \frac{p_i}{q_i}$. Correlations between two different random variables X and Y are measured by the mutual information, $I(X:Y) = H(X)+H(Y)-H(X,Y)$, where $H(X,Y) = -\sum_{i,j} p_{ij} \log p_{ij}$ is the joint entropy and p_{ij} is the probability of outcomes x_i and y_j both occurring. The mutual information measures how much information X and Y have in common. It may also be defined as a special case of the relative entropy, since it is a measure of how distinguishable a joint probability distribution p_{ij} is from a completely uncorrelated pair of distributions $p_i p_j$, $H(p_{ij}||p_i p_j) = H(p_i) + H(p_j) - H(p_{ij})$. In the quantum context, the results of a projective measurement $\{E_y\}$ on a state represented by a density matrix, ρ, comprise a probability distribution $p_y = Tr(E_y \rho)$. Von Neumann showed that the highest entropy of any of these probability distributions generated from the state ρ was achieved by the probability distribution composed of the eigenvalues of the state, $\lambda = \{\lambda_i\}$ [3]. This probability distribution would arise from a projective measurement onto the state's eigenvectors. The Von Neumann entropy is then defined as $S(\rho) = -Tr(\rho \log \rho) = H(\lambda)$. The classical relative entropy and classical mutual information also have analogues in the quantum domain. The quantum relative entropy of a state ρ with respect to another state σ is defined as $S(\rho||\sigma) = -S(\rho) - Tr(\rho \log \sigma)$. The joint entropy $S(\rho_{AB})$ for a composite system ρ_{AB} with two subsystems A and B is given by $S(\rho_{AB}) = -Tr(\rho_{AB} \log \rho_{AB})$ and the Von Neumann mutual information between the two subsystems is defined as $I(\rho_{AB}) = S(\rho_A) + S(\rho_B) - S(\rho_{AB})$. As in the classical case, the mutual information is the relative entropy between ρ_{AB} and $\rho_A \otimes \rho_B$. The mutual information is usually used to measure the total correlations between the two subsystems of a bipartite quantum system. The entanglement of a bipartite quantum state ρ_{AB} may be measured by how distinguishable it is from the 'nearest' separable state, as measured by the relative entropy. Relative entropy of entanglement, defined as

$$E_{RE}(\rho_{AB}) = \min_{\sigma_{AB} \in D} S(\rho_{AB}||\sigma_{AB})$$

has been shown to be a useful measure of entanglement (D is the set of all separable or disentangled states) [5]. Note that $E_{RE}(\rho_{AB}) \leq I(\rho_{AB})$, by definition of $E_{RE}(\rho_{AB})$, since the mutual information is also the relative entropy between ρ_{AB} and a completely disentangled state, $I = S(\rho_{AB}||\rho_A \otimes \rho_B)$ and so must be higher than the minimum over all disentangled states. There are many other ways of measuring the entanglement of a bipartite quantum state [9], but they could all be unified under the formalism of relative entropy [6]. Another advantage of relative entropy is that it can be generalized to any number of subsystems, a property that will be very useful in understanding the measurement process when the environment is also present.

There has been some work on the more general problem of splitting information in a particular quantum state into a classical and a quantum part [8,11,12]. Consider performing a general measurement on the state, ρ, given by $A_i^\dagger A_i$, such that $\rho^i = \frac{A_i^\dagger \rho A_i}{tr(A_i^\dagger \rho A_i)}$. The final state after the measurement is then $\sum_i A_i^\dagger \rho A_i = \sum_i p_i \rho^i$. The entropy of the residual states is $\sum_i p_i S(\rho^i)$. The classical information obtained by measuring outcomes i with probabilities p_i is $H(p)$. If the states ρ^i have support on orthogonal subspaces, then the entropy of the final state is the sum of the residual entropy and the classical information $S(\sum_i p_i \rho^i) = H(p) + \sum_i p_i S(\rho^i)$. Recently we have suggested that correlations in a state ρ_{AB} can also be split into two parts, the quantum and the classical part [8]. The classical part is seen as the amount of information about one subsystem, say A, that can be obtained by performing a measurement on the other subsystem, B. The resulting measure is the difference between the initial and the residual entropy [8]:

$$C_B(\rho_{AB}) = \max_{B_i^\dagger B_i} S(\rho_A) - \sum_i p_i S(\rho_A^i) \tag{1}$$

where $B_i^\dagger B_i$ is a Positive Operator Valued Measure performed on the subsystem B and $\rho_A^i = tr_A(B_i \rho_{AB} B_i^\dagger)/tr_{AB}(B_i \rho_{AB} B_i^\dagger)$ is the remaining state of A after obtaining the outcome i on B. Alternatively, $C_A(\rho_{AB}) = \max_{A_i^\dagger A_i} S(\rho_B) - \sum_i p_i S(\rho_B^i)$ if the measurement is performed on subsystem A instead of on B. Clearly $C_A(\rho_{AB}) = C_B(\rho_{AB})$ for all states ρ_{AB} such that $S(\rho_A) = S(\rho_B)$ (e.g. pure states). It remains an open question whether this is true in general (this will not affect our measurement analysis as there the apparatus is always measured to infer the state of the system and never the other way round). This measure is a natural generalisation of the classical mutual information, which is the difference in uncertainty about the subsystem B (A) before and after a measurement on the correlated subsystem A (B). Similarly, Eq. (1) represents the difference in Von Neumann entropy before and after the measurement. Note the similarity of the definition to the Holevo bound which measures the capacity of quantum states for classical communication [14]. The following example provides an illustration and will be the key to our discussion of the quantum measurement. Consider a bipartite separable state of the form

$\rho_{AB} = \sum_i p_i \left|i\right\rangle \left\langle i\right|_A \otimes \rho_B^i$, where $\{|i\rangle\}$ are orthonormal states of subsystem A. Clearly the entanglement of this state is zero. The best measurement that Alice can make to gain information about Bob's subsystem is a projective measurement onto the states $\{|i\rangle\}$ of subsystem A. Therefore the classical correlations are given by

$$C_A(\rho_{AB}) = S(\rho_B) - \sum_i p_i S(\rho_B^i)$$

For this state, the mutual information is also given by

$$I(\rho_{AB}) = S(\rho_B) - \sum_i p_i S(\rho_B^i)$$

This is to be expected since there are no entangled correlations and so the total correlations between A and B should be equal to the classical correlations. This measure of classical correlations has other important properties such as $C(\rho_{AB}) = 0$ if and only if $\rho_{AB} = \rho_A \otimes \rho_B$ and the fact that it is invariant under local unitary transformations. The key property satisfied by this measure of classical correlations is that it is non-increasing under any general local operations as shown in [8] which justifies its name as a measure of classical correlations.

Let us now introduce the general framework for a quantum measurement (for a special case see [13]). We have a system in the state $|\Psi\rangle = \sum_i a_i |i\rangle$, and an apparatus in the state $\rho = \sum_i r_i |r_i\rangle\langle r_i|$ in the eigen-basis. The purpose of a measurement is to correlate the system with the apparatus so that we can extract the information about the state $|j\rangle$ of the system. In a perfect measurement, by looking at (measuring) the apparatus we can unambiguously identify the state of the system. Therefore, when the system is in the state $|j\rangle$ we would like the apparatus to be in the state ρ_j, such that $\rho_i\rho_j = 0$, i.e. different states of the apparatus lie in orthogonal subspaces and can be discriminated with a unit efficiency. If this condition is not fulfilled, which is in general (i.e. in reality) the case, then the measurement is imperfect and the amount of information obtained is not maximal (this is what defines an "imperfect measurement"). We now compute the amount of information gained in general and show that it is more appropriately identified with the classical rather than quantum correlations (entanglement) between the system and the apparatus. Suppose that the measurement transformation is given by a unitary operator, U, acting on both the system and the apparatus, such that

$$U(\rho \otimes |i\rangle\langle j|)U^\dagger = \rho_{ij} \otimes |i\rangle\langle j|$$

where we assume that the measurement transformation acts such that the state $|r_k\rangle|l\rangle$ of the apparatus and the system respectively is transformed into the state $|\tilde{r}_{kl}\rangle|l\rangle$, such that the states of the apparatus corresponding to different system states are orthogonal $\langle\tilde{r}_{ij}|\tilde{r}_{ik}\rangle = \delta_{jk}$. We see that the measurement is such that the new apparatus state depends on the state of the system. This is exactly how correlations between the two are established. Then, the initial state is transformed into

$$\rho_f = \sum_{ij} a_i a_j^* \rho_{ij} \otimes |i\rangle\langle j| = \sum_i |a_i|^2 \rho_{ii} \otimes |i\rangle\langle i| + \sum_{i\neq j} a_i a_j^* \rho_{ij} \otimes |i\rangle\langle j|$$

The first term on the right hand side indicates how much information this measurement carries. We will now measure the apparatus and try to distinguish the states ρ_{ii} to the best of our ability. Once we confirm that the apparatus is in the state ρ_{jj}, then we can infer that the system is in the state $|j\rangle$. The amount of information about the state of the apparatus (and hence the state of the system), I_m, is given by the well-known Holevo bound [14]:

$$I_m = S(\sum_i |a_i|^2 \rho_{ii}) - \sum_i |a_i|^2 S(\rho_{ii}) \tag{2}$$

As we have seen, this quantity is also equal to the amount of classical correlations between the system and the apparatus in the state $\rho'_f = \sum_i |a_i|^2 \rho_{ii} \otimes |i\rangle\langle i|$, which is, in this case, the same as the von Neumann mutual information between the two. Note that this state is only classically correlated and there is no entanglement involved. The amount of entanglement in the state ρ_f, on the other hand, will in general be non-zero. This may be difficult to calculate, however, we can provide lower and upper bounds. The lower bound on the entanglement between the system and the apparatus is

$$E(\rho_f) \geq S(\sum_i |a_i|^2 \rho_{ii}) - S(\rho_f) = S(\sum_i |a_i|^2 \rho_{ii}) - S(\rho) = I_m \tag{3}$$

Therefore, the entanglement between the system and the apparatus is larger than or equal to the classical correlations between the two which quantify the amount of information that the measurement carries. So, this shows that the information in a quantum measurement is correctly identified with the classical correlations between the apparatus and the system rather than the entanglement or the mutual information between the two in the final state, ρ_f. Only in the limiting case of the pure apparatus (analyzed first by von Neumann [3] and then used by Everett [1] in his interpretation of quantum mechanics) do we have that the amount of information in the measurement is equal to the entanglement, which becomes the same as the classical correlations, while the sum of the quantum and classical correlations is then equal to the mutual information in the state.

We can recast this relationship in the form of an "uncertainty relation" between the initial mixedness of the apparatus and the amount of information gained. So, from the fact that $I_m = S(\sum_i |a_i|^2 \rho_{ii}) - S(\rho)$, we have that

$$I_m + S(\rho) = S(\sum_i |a_i|^2 \rho_{ii}) \leq \log N \tag{4}$$

where N is the dimension of the apparatus. Thus we see that the sum of the initial mixedness of the apparatus and the amount of information the measurement obtains is always smaller than a given fixed value: the larger $S(\rho)$, the smaller I_m. When ρ is maximally mixed (and therefore $S(\rho) = \log N$), then no information can be extracted from the measurement. Note that this relation is different to the usual "information versus disturbance" law in a quantum measurement as well as to the usual entropic uncertainty relations of incompatible observables. Every measurement that extracts information from a quantum system also disturbs the state, and without this disturbance there would be no information gain possible. The initial state of the system in our above scenario was $\sum_i a_i|i\rangle$, while the final state is a mixture of the form $\sum_i |a_i|^2 |i\rangle\langle i|$. The disturbance to the state can be measured as a distance between the final and the initial state. We choose the relative entropy to quantify this difference. So, while the information in the measurement is given by $S(\sum_i |a_i|^2 \rho_{ii})$, the disturbance is

$$D = S(|\Psi\rangle\langle\Psi| || \sum_i |a_i|^2 |i\rangle\langle i|) = -\sum_i |a_i|^2 \log |a_i|^2$$

which is the same as the maximum amount of information possible from this measurement. So, the measurement described above always maximally disturbs the state, and the reason why this does not lead to the maximum information gain is because the apparatus state is mixed. The system could be disturbed less by adjusting the overlap between the states of the apparatus $|\tilde{r}_{ij}\rangle$, so that they are not orthogonal to each other. In general we can require that $\langle \tilde{r}_{ij}|\tilde{r}_{ik}\rangle = a_{jk}$, such that a_{jk} is a complex number with less than unit modulus. We will not treat this case here: it is mathematically more demanding, but does not illuminate the measurement issue any better.

In order to show that some form of entanglement is still important (albeit not the one between the system and the apparatus) we revisit the same measurement scenario, but from the "higher Hilbert space perspective". This is done by adding the environment to the apparatus so that the joint state is pure, $|\Psi_{EA}\rangle$. We briefly note that our treatment differs form the usual "environment induced collapse" and decoherence as in, for example, [15,16]. In our case, the environment is not there to cause the disappearence of entanglement between the system and the apparatus, but is there to purify the initally mixed state of the apparatus. We do not have any state reduction or collapse, but just different ways in which we can express the classical correlations between the system and the apparatus. The measurement transformation is now given by

$$|\Psi_{EA}\rangle \otimes \sum_i a_i|i\rangle \longrightarrow \sum_i a_i|\Psi^i_{EA}\rangle|i\rangle$$

where $|\Psi_{EA}\rangle = \sum_i \sqrt{r_i}|e_i\rangle|r_i\rangle$ and $|e_i\rangle$ is an orthonormal basis for the environmental states. We see that when the environment is traced out, the state of the apparatus is equal to ρ. Now, the measurement implements a unitary transformation so that each of the states of the apparatus changes according to which state of the system it interacts with. Therefore we see that the i-th state of the environment and the apparatus after the interaction is given by $|\Psi^i_{EA}\rangle = \sum_j \sqrt{r_j}|e_j\rangle|\tilde{r}_{ji}\rangle$. To make a link with the first picture of the measurement, we trace out the environment to obtain:

$$\rho_{A'S'} = \sum_{ij} a_i a_j^* (\sum_k \langle e_k|\Psi^i_{EA}\rangle\langle\Psi^j_{EA}|e_k\rangle) \otimes |i\rangle\langle j|$$

and, thus, the quantity in brackets can be identified with $\rho_{ij} = \sum_k r_k |\tilde{r}_{ki}\rangle\langle\tilde{r}_{kj}|$. Therefore, since we have no access to the environment, our task is to discriminate the states ρ_{ii}, and therefore identify the corresponding states $|i\rangle$ of the system and this was done in the previous analysis. If, on the other hand, we had access to the environment, the measurement could be perfect.

We first apply entropic considerations to the "environment-apparatus-system" tripartite state. First of all, the initial and the final entropy of the environment are the same as its state remains unchanged, and this value is the same as the initial entropy of the apparatus, $S(\rho)$. As we have seen, this is an important quantity as it determines how much information can be extracted from a measurement: the more mixed the initial state of the apparatus, the less information can be extracted. If the initial state is maximally mixed (say it is a thermal state with an arbitrarily high temperature), then there can be no information gain during the measurement. The initial entropy of the apparatus is also equal to the entropy of the system and the apparatus after the measurement, $S(\rho_{A'S'}) \equiv S(\rho_f)$, as well as the amount of entanglement between the environment and the system and the apparatus together, $E_{E:(A'S')}$, after the measurement. The entanglement and the mutual information between the environment and the apparatus after the measurement are always less than or equal to their value before the measurement (since the systems becomes correlated to the apparatus during the measurement).

We now use the recently derived three-party entanglement bounds to provide further constraints on the measurement. For any pure tripartite states σ_{ABC} we have that [7]:

$$\begin{aligned}&\max\{E(\sigma_{AB})+S(\sigma_C),E(\sigma_{AC})+S(\sigma_B),E(\sigma_{BC})+S(\sigma_A)\}\leq E(\sigma_{ABC})\\&\leq\min\{S(\sigma_A)+S(\sigma_B),S(\sigma_A)+S(\sigma_C),S(\sigma_B)+S(\sigma_C)\}\end{aligned} \tag{5}$$

Applying this to our measurement scenario we obtain that

$$S(\rho)\leq E_{E:A':S'}-E_{A':S'}$$

where the subscripts E, A, S indicate the environment, the apparatus and the system respectively. The primes on the subscripts indicate states after the measurement. We see that the closer the tripartite entanglement to the entanglement between the system and the apparatus (with the environment disentangled), the more efficient the measurement. We immediately conclude that the necessary condition for the equality between the two entanglements is that the initial entropy of the apparatus is zero. It should be remembered, however, that the measurement can still be perfect even though the apparatus is not pure and this is because the relevant quantity is the classical correlations between the system and the apparatus and not their entanglement. In that context we can also derive from the inequality in (5) that $E_{E:A'}+S(\rho_{S'})\leq E_{E:A':S'}$, so that

$$E_{E:A'}+I_m\leq E_{E:A':S'}-S(\rho)\leq S(\rho'_A)$$

Thus, the sum of the information form the measurement and the final entanglement between the environment and the apparatus is limited by the final entropy of the apparatus and therefore by $\log N$. Again we see that the larger the information we want, the smaller the entanglement with the environment and the apparatus will be. So, in fact, for the measurement to be efficient we wish the environment *not* to become entangled with the apparatus to a large extent (after the measurement). We should mention at the end that our example is somewhat simplified in that the environment will not, in reality, be passive throughout the process. It would instead interact with both the system and the apparatus making the measurement even less effective, although all the above results would still apply.

In this letter we have analyzed the information gained in a quantum measurement when the apparatus used to extract this information is initially in a mixed state (or entangled to its environment). This is a realistic scenario as the apparatus is usually assumed to be macroscopic and it is consequently in thermal equilibrium with its own environment. We have shown that the amount of information is correctly identified with the amount of classical correlations between the system and the apparatus after their correlation is established and derived an entropic uncertainty relation between this amount and the mixedness of the initial state. Further light on quantum measurement was then shed by purifying the apparatus and including its own environment in the analysis and we have shown that the entanglement between the environment and the apparatus plays a role in limiting the information gain in the measurement. Among open problems highlighted by this work are to extend the analysis to non-orthogonal states of the apparatus in the measurement transformation and to prove that the information gain is symmetric between the system and the apparatus.

Acknowledgments. This work is funded by Engineering and Physical Sciences Research Council, the European grant EQUIP and Hewlett-Packard. The great hospitality of the staff at the Maternity Ward of the Raigmore Hospital in Inverness, Scotland is acknowledged where part of this work was completed.

---

[1] H. Everett, Rev. Mod. Phys. 29, 454 (1957); see also H. Everett, editors B. S. DeWitt, R. Neill Graham eds., *The many-worlds Interpretation of Quantum Mechanics*, (Princeton Series in Physics, Princeton University Press (1973)).

[2] J. A. Wheeler, W. H. Zurek eds., *Quantum Theory and Measurement*, (Princeton Series in Physics, Princeton University Press (1983)).
[3] J. von Neumann, *Mathematical Foundations of Quantum Mechanics*, (Princeton University Press, New York, 1955).
[4] G. Lindblad, Comm. Math. Phys. **33** 305 (1973).
[5] V. Vedral and M. B. Plenio, Phys. Rev. A, **57**, 1619 (1998).
[6] L. Henderson and V. Vedral, Phys. Rev. Lett. **84**, 2263 (2000).
[7] M. B. Plenio and V. Vedral, J. Phys. A **34**, 6997 (2001).
[8] L. Henderson and V. Vedral, J. Phys. A **34**, 6899 (2001).
[9] V. Vedral, Rev. Mod. Phys. **74**, 297 (2002).
[10] E. Shannon and W. Weaver, *The Mathematical Theory of Communication*, (University of Illinois Press, Urbana, IL (1949).
[11] H. Ollivier and W. H. Zurek, Phys. Rev. Lett. **88**, 017901 (2002).
[12] B. M. Terhal, M. Horodecki, D. W. Leung and D. P. DiVincenzo, lanl e-print no. 0202044 (2002).
[13] S. Bose and V. Vedral, lanl e-print no. 0004016 (2000).
[14] A.S. Holevo, Probl. Peredachi Info. **9**, 177 (1973).
[15] W. H. Zurek, Phys. Rev. D **24**, 1516 (1981).
[16] M. H. Partovi, Phys. Lett. A **137**, 445 (1985).